Big Data and Computing Visions

www.bidacv.com

Big. Data. Comp. Vis. Vol. 5, No. 2 (2025) 102–118.

Paper Type: Original Article

Sentiment Analysis for Identifying Depression through Social Media Texts Using Machine Learning Technique

Barige Bharadwaj[1], Sukanta Nayak[2], Paresh Kumar Panigrahi[2,*]

[1] School of Advanced Sciences, VIT-AP University, Amaravati, Andhra Pradesh, India; barigebharadwaj@gmail.com.

[2] Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati, AP, India; sukantgacr@gmail.com; jrpareshkp@gmail.com.

Citation:

Received: 11 September 2024 *Revised: 24 November 2024* *Accepted: 11 January 2025*	Bharadwaj, B., Nayak, S., & Panigrahi, P. K. (2025). Sentiment analysis for identifying depression through social media texts using machine learning technique. *Big data and computing visions, 5*(2), 102-118.

Abstract

This paper presents a detailed exploration of the evolving landscape of depression detection through Sentiment Analysis (SA) in online communication platforms. With depression being a widespread and often undetected mental health concern, leveraging technology for early intervention is crucial. The study delves into three key approaches: lexicon-based methods, machine learning algorithms, and hybrid models, providing a thorough analysis of their strengths and limitations. It traces the historical evolution of SA, highlighting pivotal advancements, including deep learning techniques and multimodal data integration. The paper emphasizes the challenges, such as privacy concerns and algorithmic biases, and proposes future research directions, emphasizing multi-lingual analysis and interdisciplinary collaboration. The findings underscore the transformative potential of SA in reshaping mental health interventions and fostering inclusivity in support systems. Depression is a widespread challenge, often difficult to detect and monitor effectively. This paper explores how we can better understand and support individuals experiencing depression through SA. We delve into various methods used to analyze the emotions expressed in text, speech, and behaviour to identify signs of depression. We focus on the importance of spotting these signs early, assessing risks, and tailoring support for each person. Moreover, we discuss how we can advance these methods to improve mental health care. By looking closely at the current methods and their practical use, we aim to shed light on SA's role in caring for mental health. The goal is to emphasize the need for ongoing research and innovation to make these analyses even more effective in monitoring and supporting individuals dealing with depression.

Keywords: Depression detection, Sentiment analysis, Lexicon-based methods, Machine learning algorithms, Hybrid models.

1 | Introduction

Depressive disorder, commonly known as depression, is a pervasive mental health condition characterized by persistent feelings of sadness, loss of interest or pleasure in activities, and alterations in daily functioning. It significantly impacts an individual's relationships, work, and overall quality of life.

Corresponding Author: jrpareshkp@gmail.com

https://doi.org/10.22105/bdcv.2025.499270.1239

According to the World Health Organization (WHO), depression isn't merely a fluctuation in mood but a condition that can manifest in various forms, affecting individuals regardless of age, gender, or social status. Stressful life events, trauma, or significant losses can contribute to its development, with women more likely to be affected than men.

Citing the WHO's March 2023 statement, it's estimated that approximately 3.8% of the global population experiences depression. These rates are notably higher in specific demographics, such as among adults, with 5% affected (4% of men and 6% of women) and even higher among individuals over 60 years old, accounting for about 5.7% of this population. The impact on women, as the statistics indicate, is significantly greater, and the rates soar during pregnancy and postpartum periods, affecting more than 10% of women globally.

The severity of depression is starkly outlined by the fact that more than 700,000 people lose their lives due to suicide every year, making it the fourth leading cause of death among 15–29-year-olds. However, despite effective treatments available for mental disorders, a staggering 75% of individuals in low- and middle-income countries do not receive the necessary mental health care. This treatment gap is largely due to various barriers, including inadequate investment in mental health care, a shortage of trained healthcare professionals, and the pervasive social stigma associated with mental health conditions.

This WHO statement from March 2023 underscores the critical need for increased awareness, resources, and improved mental health care infrastructure to address the global burden of depression. Depression, as highlighted by the WHO, is a pervasive and debilitating mental health condition affecting millions worldwide. Its impact reaches far beyond an individual's mental state, significantly influencing daily life and relationships and even contributing to global mortality rates, particularly due to suicide.

Given the gravity and prevalence of depression, understanding and analyzing sentiments and emotions play a crucial role in identifying, addressing, and potentially preventing such mental health disorders. This need has spurred the development and application of Sentiment Analysis (SA), a branch of Natural Language Processing (NLP) that involves identifying, extracting, and interpreting emotions, opinions, and attitudes expressed in textual data.

SA, also known as opinion mining, has garnered increasing attention across various domains, including mental health research, to comprehend sentiments embedded within textual data. The application of SA in mental health has the potential to provide insights into emotional states, sentiment shifts, and indicators of depressive tendencies, offering an additional dimension to understand and support individuals experiencing such conditions.

This paper aims to explore the landscape of existing research on SA in mental health, particularly focusing on its applications and limitations within the domain of depression. The work aspires not only to review and analyze existing methodologies but also to identify the gaps and potential areas for future research, acknowledging the critical need for advanced techniques and tools in the realm of mental health and SA.

By blending insights from mental health data and the capabilities of SA, this paper seeks to contribute to the advancement of methodologies that may better recognize, diagnose, and ultimately support individuals struggling with depressive disorders. This transition from discussing depression to introducing SA sets the stage for the paper's focus on exploring SA in the context of mental health, particularly depression.

2 | Advancements in Depression Detection and Sentiment Analysis

Depression is a pervasive mental health issue that significantly affects individuals' well-being and quality of life. With the increasing prevalence of depression, there is a growing need for efficient and scalable methods to detect and monitor depressive symptoms. SA, a NLP technique, has been widely explored for its potential to detect and monitor depression by analyzing text data from various sources, including social media platforms and text messages. This section provides an overview of the current research landscape in depression detection using SA, encompassing both the previously mentioned works and additional relevant studies.

Huang et al. [2] employed an attention-emotion-enhanced convolutional Long Short-Term Memory (LSTM) model to achieve high sentiment categorization performance. However, their model's inability to balance ideal parameters resulted in reduced performance and robustness. Basiri et al. [3] conducted a study highlighting the attraction of coronaviruses to people from different countries at various times and severity levels, although their data retrieval did not contribute to keyword analysis for tweets. Gandhi et al. [4] leveraged convolutional neural networks and LSTM models to eliminate frequently occurring terms in Twitter data, although they recognized the need for improvements in feature learning methods. Jain et al. [5] used a CNN-LSTM model for unstructured user review analysis, focusing exclusively on information from an online platform. Kaur et al. [6] proposed a Hybrid Heterogeneous Support Vector Machine (H-SVM) approach, extracting tweets from various languages. However, their data collection was limited to Twitter, and traditional vector representation methods were used.

Pathak et al. [7] adopted an LSTM-based sentiment scrutiny approach that can scale with streaming short-text input and support online responses. Yet, their model couldn't effectively capture the nature of distributed representations. Basiri et al. [8] proposed an attention-based bidirectional CNN-Recurrent Neural Networks (RNN) deep model, which removed restricted features and extensive dependencies but had limited data for tasks. Wang et al. [9] introduced a Relational Graph Attention Network (R-GAT) model that facilitated batch and parallel processes but failed to account for dependency relations. Zhang et al. [10] presented the Broad Multitask Transformer Network (BMT-Net), which consistently produced positive experimental outcomes, even though unstructured data evaluations were less accurate. Dashtipour et al. [11] developed a Persian sentiment approach characterized by less over-fitting and better generalization without introducing new layers.

These studies provide insight into the evolving landscape of depression detection using SA. While they showcase various methods and approaches, the field continually evolves, exploring novel techniques and approaches to improve the accuracy and efficiency of depression detection. We expand this literature review by incorporating additional research works to offer a more comprehensive view.

Sabri and Mohamad [12] presented a method for detecting suicidal tweets using the Naïve Bayes algorithm. Their approach focused on identifying tweets with suicidal sentiments. Samareh et al. [13] integrated computer vision, signal processing, and SA to detect depression from communication data. They used a multimodal approach to improve the accuracy of depression detection. Researchers [14-16] investigated the relationship between text message sentiment and self-reported depression, shedding light on the use of self-reported data as a valuable source for analysis. Mustafa et al. [17] introduced a multiclass depression detection system based on SA, enhancing the classification of different depression levels in social media data. Ushashree et al. [18] proposed a methodology using SA to detect depression in social media users, outlining an approach to leverage sentiment features.

Furthermore, Drus and Khalid [19] conducted a systematic literature review on SA in social media, providing an overview of the existing approaches and their applications. Suman et al. [20] developed a novel SA engine for preliminary depression status estimation on social media, offering a unique approach to estimating the initial state of depression. Jo et al. [21] investigated the relationship between computational models and psychological states, aiming to identify what SA seeks to find in the context of psychological well-being. Islam et al. [22]explored depression detection using the K-Nearest Neighbours (KNN) classification technique, emphasizing the importance of choosing appropriate classification algorithms. Finally, Li et al. [23] analyzed linguistic features to detect depression stigma on social media, highlighting the significance of language patterns in detecting social stigma associated with depression.

In conclusion, the body of research on depression detection through SA is extensive and continually evolving. These studies collectively contribute to advancing computational methods for identifying and monitoring depressive symptoms across various data sources. Their diverse methodologies, strengths, and limitations reflect the ongoing efforts to develop accurate and scalable tools for addressing the critical issue of depression.

As the field matures, integrating these diverse approaches may lead to more robust and effective depression detection and monitoring systems.

3 | Overview of Step-by-Step Process of Sentiment Analysis

3.1 | Data Collection

Gather text data from various sources (e.g., social media, forums). For depression detection via SA, data collection from social media platforms involves leveraging APIs to access and retrieve public data, specifically user-generated content like posts and comments. This process targets expressions reflecting emotional states within the textual information. However, it necessitates respecting privacy and ethical considerations by anonymizing data and adhering to platform policies. Overall, this approach harvests valuable insights from social media to understand and analyze sentiments linked to depression. The dataset, denoted as $\{D = d_1, d_2, d_3, \dots\dots. d_n\}$ consists of $'n'$ text data points.

3.2 | Preprocessing

Cleaning and Tokenization: each d_i is cleaned by removing non-essential elements like URLs and punctuation. After cleaning, the data is tokenized into individual words or phrases, forming a token set for each d_i denoted as $T = \{t_1, t_2, \dots, t_m\}$.

Standardization: all words are converted to lowercase to standardize the text data.

Stop words removal: eliminate common words (stop words) that may not add significant meaning to the text data.

3.3 | Lemmatization and Stemming

Apply lemmatization or stemming to reduce words to their root forms. Lemmatization involves converting words to their base or dictionary forms, while stemming reduces words to their word stems. This step helps reduce the text data's dimensionality, combining various forms of a word into a single representation.

3.4 | Feature Extraction

Bag-of-Words (BoW) and TF-IDF: Create a numerical representation of text data, either through a BoW model or by applying the TF-IDF technique to create matrices, e.g., M_{BoW}, M_{TF-IDF} that indicate word occurrences and importance.

3.5 | Sentiment Analysis Algorithms

I. Lexicon-based analysis: assign sentiment scores to words using predefined dictionaries that map words to their emotional intensities.

II. Machine learning models (e.g., SVM, naive bayes): classify text data into sentiment categories using features derived from the numerical representation.

III. Deep learning models (e.g., LSTM, CNN): utilize neural networks to capture intricate patterns and relationships in text, learning representations suitable for sentiment classification.

3.6 | Sentiment Classification and Polarity Range

Classify the SA output into positive, negative, or neutral categories using class labels (e.g., 0, 1, 2) based on the models' predictions. Sentiment classification polarity assessment: assign sentiment polarity to text data based on the analysis results. Sentiments are often quantified on a numerical scale, where the range can denote various emotional intensities (e.g., negative, neutral, positive) or be more nuanced, particularly for depression detection. Polarity can be represented by real values within a specific range, commonly from -1 to 1.

3.7 | Mathematical Representation of Sentiment Polarity

The assigned Polarity (P) for a piece of text data (d_i) is represented as a continuous value. For instance, (P) may vary between -1 (highly negative sentiment) to 1 (highly positive sentiment) on a real number line. Mathematical operations like averaging the polarity scores can provide an aggregate sentiment score for a set of text data, denoted as

$$\text{Average P}=\sum_{i=1}^{n} P_i,$$

aiding in understanding the overall emotional tone in a dataset.

3.8 | Validation Metrics and Mathematical Evaluation

Evaluation of SA models: this involves employing mathematical metrics to validate the performance of SA models:

I. Confusion matrix: utilize a square matrix to describe the performance of a classification model.

II. Accuracy: mathematically expresses the ratio of correctly predicted sentiment to the total sentiment. It is computed as

$$\text{Acc} = \frac{(\text{TP} + \text{TN})}{\text{TP} + \text{TN} + \text{FN} + \text{FP}},$$

where TP denotes True Positives, TN denotes True Negatives, FP denotes False Positives, and FN denotes false negatives.

Precision: mathematical calculation of precision is given by

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}}.$$

Denoting the ratio of correctly predicted positive observations to the total predicted positives.

Recall sensitivity: represented mathematically as

$$\text{Recall} = \frac{\text{TP}}{(\text{TP}+\text{FN})}.$$

Calculating the proportion of actual positives that were correctly identified.

3.9 | Interpretation and Reporting

Analyze the sentiment output and gain insights into the presence and distribution of depressive sentiments in the text data.

3.10 | Feedback Loop and Model Refinement

Incorporate ongoing feedback to improve the model by adjusting for limitations and leveraging newly available data for iterative model enhancement.

4 | Sentiment Analysis Methods for Depression Monitoring

Depression monitoring through SA involves extracting emotional cues from text data, enabling the identification of individuals at risk. This section delves into three key methods: the lexicon-based approach, machine learning algorithms, and the hybrid approach. Each method offers distinct advantages and challenges in detecting depression-related sentiment patterns within online communication platforms.

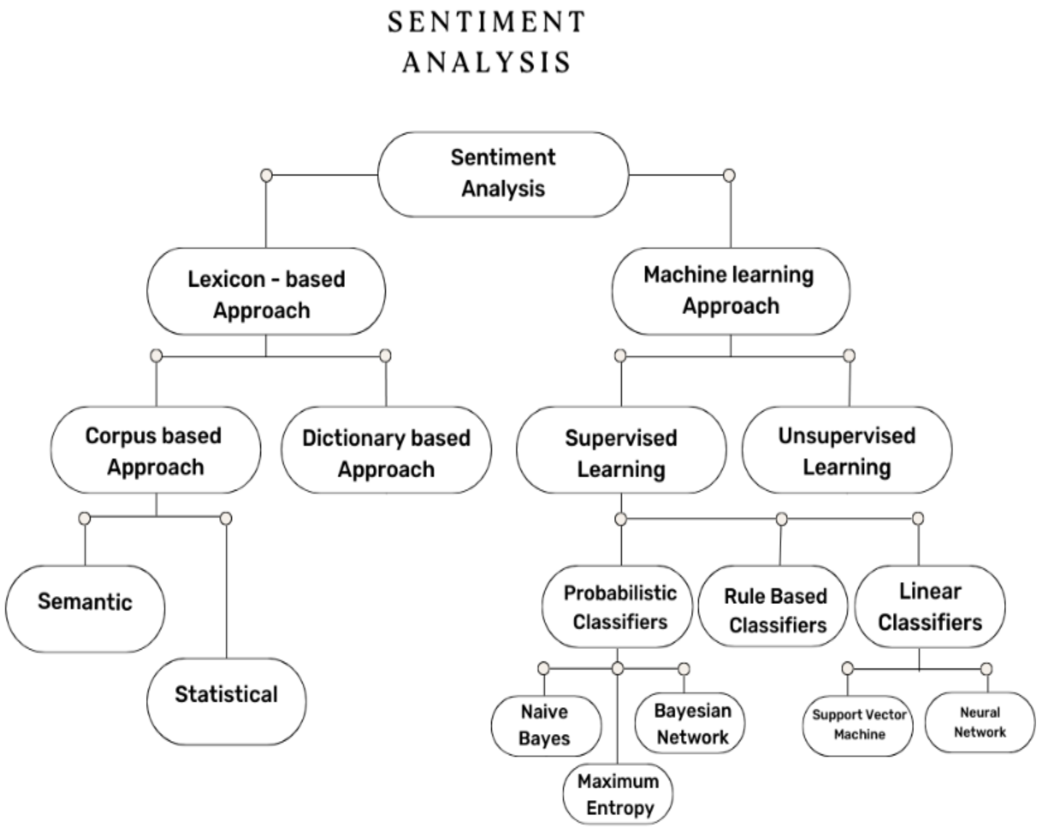

Fig. 1. Flow chart of the sentimental analysis.

4.1|Lexicon-based Approach for Depression Monitoring

In the context of depression monitoring, the Lexicon-Based Approach presents an initial method to gauge sentiment in textual data from online communication platforms. This approach seeks to identify and assess depressive content based on predefined sentiment lexicons, which associate words with sentiment labels. While simplistic, this technique can provide valuable insights into the emotional state of individuals using such platforms.

4.1.1 |Valence aware dictionary and sentiment reasoner

Purpose: Valence Aware Dictionary And Sentiment Reasoner (VADER) is a lexicon and rule-based SA tool that quantifies the sentiment of a piece of text into a single numerical score.

I. Algorithm**:** the VADER model uses a dictionary of words lexicon with predefined scores representing the positivity or negativity of terms and a set of rules to determine the sentiment of a text. These scores are then combined using a specific formula to generate a composite score.

II. Advantages**:** VADER can analyze both individual words and idiomatic phrases, capturing the context and intensity of emotions in text. It's particularly useful for social media text analysis.

III. Disadvantages**:** while VADER is effective for short informal texts, its performance might vary in other domains. It may not handle sarcasm or complex sentences effectively.

IV. Mathematical Formula: the compound score is calculated as the sum of the valence scores of each word in the text, normalized between -1 and 1.

$$\mathrm{C} = \frac{\sum \text{valence scores}}{\mathrm{N} \times \text{maximum valence scores}},$$

where C = The compound score, valence scores = Individual valence scores of words, N = Total number of words in the text, and maximum valence score = the maximum possible valence score.

4.1.2 | Sent wordnet and text blob

I. Purpose: sent WordNet and Text Blob are also lexicon-based models that assign pre-calculated sentiment scores to words and compute an overall sentiment score for a text based on these word scores.

II. Algorithm: sent WordNet assigns scores based on word meanings in the word net database. Text blob utilizes a predefined set of scores for words and rules for sentence structure to gauge sentiment.

III. Advantages: both models offer a quick and simple way to gauge sentiment from text by considering the pre-assigned scores of words in a given corpus.

IV. Disadvantages: these models might lack the ability to capture context effectively and struggle with nuanced or ambiguous text expressions.

Mathematical formula

Unlike VADER, these models might not have a single comprehensive mathematical equation for sentiment score computation; they rely on word scores combined through various linguistic rules to compute overall sentiment. Each model adopts a distinct approach to analyzing sentiment in text, some relying on pre-calculated scores associated with words or phrases and others using a rule-based system to gauge sentiment from the composition of individual terms within the text.

Mechanism

I. Data preprocessing and cleaning: the text data collected from online communication platforms, such as social media and messaging apps, undergoes a preprocessing phase to remove extraneous elements like special characters and stopwords. This ensures that the subsequent analysis focuses solely on relevant words.

II. Tokenization: after preprocessing, the text is tokenized, breaking it down into individual words or n-grams. This step enables SA in smaller units, aiding the classification process.

III. Sentiment lexicon matching: the heart of the lexicon-based approach lies in matching the tokens from the text against a sentiment lexicon tailored to depression-related sentiments. These lexicons are curated to include words and phrases indicative of depressive emotions, allowing for sentiment inference.

IV. Sentiment score calculation: once matching words are identified, sentiment scores are assigned based on their emotional connotations. Words denoting positivity contribute positively to the sentiment score, whereas negative words contribute negatively. The overall sentiment score emerges from the cumulative effect of these individual scores.

V. Aggregation and classification: aggregating the sentiment scores yields an overall sentiment score for the text. This score aids in classifying the text as either positive, negative (indicative of depressive sentiment), or neutral. Predefined thresholds determine the classification outcome.

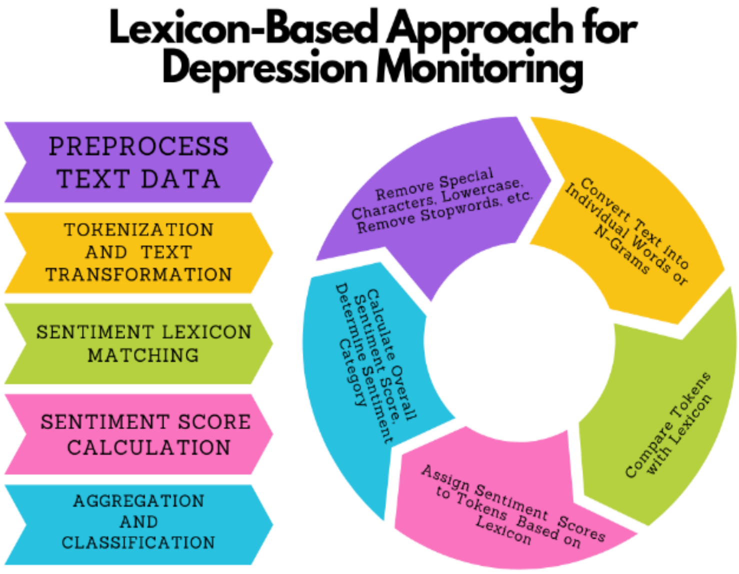

Fig. 2. Flow chart for lexicon-based approach.

I. Advantages: the lexicon-based approach's simplicity and speed make it suitable for quick sentiment assessment of large datasets. However, its performance can be hindered by nuanced language use and contextual complexities, impacting accuracy.

II. Rapid insights: the lexicon-based approach swiftly provides initial insights into sentiment patterns in online communication related to depression.

III. Ease of implementation: this method's simplicity makes it accessible, enabling researchers and practitioners to assess sentiment quickly.

IV. Transparency: the sentiment lexicons offer transparency, allowing users to comprehend sentiment classifications.

Challenges

I. Subtleties in language: The lexicon-based approach may struggle with capturing subtle nuances, idiomatic expressions, and sarcasm common in online communication.

II. Contextual complexity: the approach lacks contextual understanding, potentially leading to inaccurate classifications when emotional language is used in diverse contexts.

III. Risk of misclassification: the presence of words with dual meanings, conveying both positive and negative sentiments, can lead to misclassifications.

4.2 | Machine Learning Algorithms for Depression Monitoring

In the realm of depression monitoring, machine learning algorithms offer a potent framework that capitalizes on data-driven insights. By learning patterns from labelled datasets, these algorithms can discern intricate nuances in sentiment expressions, enabling a deeper understanding of emotional states within online communication platforms. While providing detailed information about all machine learning models, including their purposes, advantages, disadvantages, and mathematical equations, requires a considerable amount of specific data, I can give you an overview of some popular machine learning models used in SA.

4.2.1 | Naive bayes for depression monitoring

Naive bayes is a probabilistic classification algorithm that computes the probability of a particular class given the input features. In the context of depression monitoring, data preprocessing and cleaning, the initial phase involves removing extraneous elements from the textual data while tokenizing it into individual units, such as words or n-grams.

I. Feature extraction: To facilitate the application of Naive Bayes, textual data is transformed into numerical feature vectors using techniques like BoW or TF-IDF. These representations capture word frequencies or importance.

II. Training: a labelled dataset containing instances classified as depressive or non-depressive is used to train the Naive Bayes classifier. During training, the algorithm calculates probabilities based on the occurrence of features in both classes.

III. Classification: the same preprocessing and feature extraction steps are applied for an unseen text. The trained Naive Bayes classifier predicts the sentiment label based on the calculated probabilities, choosing the class with the higher probability.

IV. Purpose: it's a probabilistic classifier based on Bayes' Theorem, often used for text classification tasks.

V. Advantages: simple, computationally efficient, and works well with a small amount of data.

VI. Disadvantages: relies on the assumption of independence between features.

VII. Mathematical equation: the probability of a text belonging to a specific sentiment Category (C) given the Features (F) present in the text can be calculated using Bayes' theorem as follows:

$$\mathrm{P(C/F)} = \frac{\mathrm{P(C) \times P(F/C)}}{\mathrm{P(F)}}.$$

Here, $\mathrm{P(C / F)}$ is the posterior probability of the text belonging to the sentiment category C given its features F. $\mathrm{P(C)}$ is the prior probability of the text belonging to category C. $\mathrm{P(F / C)}$ is the likelihood, which is the probability of features F occurring given the category C. $\mathrm{P(F)}$ is the probability of features F is occurring.

4.2.2 | Support vector machines for depression monitoring

SVM is a robust machine learning algorithm used for classification tasks, including SA within depression monitoring. Data preprocessing and cleaning: The initial steps remain consistent with other data preprocessing and tokenization techniques.

I. Feature extraction: similar to Naive Bayes, feature extraction involves transforming text data into numerical feature vectors using techniques like BoW or TF-IDF.

II. Training: using a labelled dataset, where instances are labelled as depressive or non-depressive, the SVM algorithm is trained. It learns to establish an optimal hyperplane that maximally separates the two classes in the feature space.

III. Classification: when confronted with new text, the preprocessing and feature extraction steps are replicated. The trained SVM model predicts the sentiment label by assessing the position of the input text relative to the learned hyperplane.

IV. Purpose: effective for binary classification, widely used in text and hypertext categorization.

V. Advantages: works well with high-dimensional data and can handle nonlinear relationships.

VI. Disadvantages: requires extensive tuning for optimal performance.

VII. Mathematical equation: the decision boundary in SVM is often represented as a hyperplane.

$$\mathrm{w} \,.\, \mathrm{x} + \mathrm{b} = 0.$$

w is the weight vector. x is the input feature vector. b is the bias or intercept. The output is determined by $\mathrm{w} \,.\, \mathrm{x} + \mathrm{b} = 0$, where a positive result implies one class and a negative result implies another.

Advantages and challenges

Machine learning algorithms excel in capturing subtle sentiment patterns and adapting to varied linguistic expressions. In the context of depression monitoring, they have the potential to unravel intricate emotional nuances that can aid in identifying individuals at risk. However, the efficacy of these algorithms hinges on the

quality and quantity of training data. The interpretation of results might be challenging in cases of highly imbalanced datasets, and overfitting can arise with insufficient data.

Conclusion

Incorporating naive bayes and Support Vector Machines (SVM) into depression monitoring widens the spectrum of insights derived from online communication platforms. These algorithms provide a data-driven approach that can harness the complexity of sentiment expressions, thereby contributing to the identification and intervention of depression in an increasingly digital age.

Flowchart

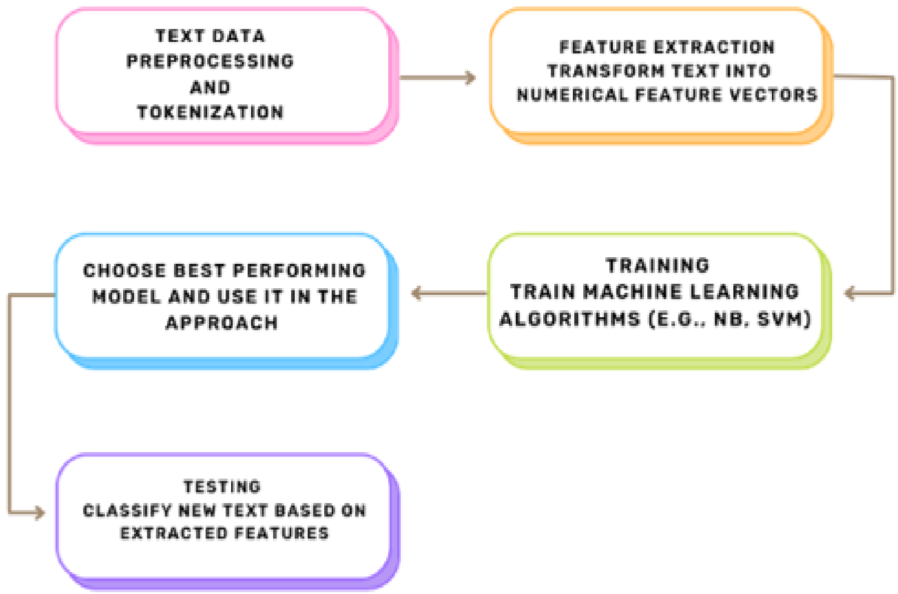

Fig. 3. Flowchart on machine learning-based approach.

While Naive Bayes and SVM are prominent models extensively employed in the realm of machine learning for depression detection, numerous other approaches have gained traction for SA. Models such as random forest, logistic regression, and gradient boosting also exhibit notable utility in processing textual data for mental health analysis. Moving beyond these conventional methods, the exploration of advanced models, particularly in the realm of deep learning, offers new avenues for more sophisticated SA. Deep learning models like RNNs and LSTM networks have displayed promising capabilities in capturing intricate linguistic nuances and sequential dependencies within text data, marking a significant shift toward more intricate, context-aware SA in depression monitoring.

4.2.3 | Recurrent neural networks for depression monitoring

I. Purpose: RNNs are designed to handle sequential data by retaining a memory of past inputs and are particularly useful for analyzing time series data, such as sequential text data in depression monitoring.

II. Algorithm: data preprocessing and cleaning: similar to other methods, RNNs require adequate data cleaning, preprocessing, and tokenization. The recurrent nature of RNNs allows the network to retain sequential information and learn from a sequence of words or characters in the text.

III. Feature extraction: RNNs inherently capture sequential dependencies, requiring minimal feature engineering. The network automatically processes the sequential data in a manner suitable for SA.

IV. Training: RNNs are trained using backpropagation through time, updating the network's weights to minimize the error in predicting the next word in the sequence based on the previous words.

V. Classification: when presented with new text, the trained RNN predicts the sentiment based on the text data's learned sequential patterns and context.

VI. Advantages: RNNs are adept at learning long-term dependencies in sequential data. They perform well in modelling context-aware SA by capturing the relationships between words or phrases in a sentence.

VII. Disadvantages: RNNs are prone to the vanishing or exploding gradient problem, making training challenging. They might struggle to capture extremely long-term dependencies.

VIII. Mathematical equation: the mathematical equation for the forward pass in a basic RNN can be represented as follows, where x represents the input, h is the hidden state, and y is the output:

$$h_t = f(W \cdot x_t + U \cdot h_{t-1} + b).$$

Here, h_t is the hidden state or output at the time step.

f is the activation function (commonly, it is a hyperbolic tangent or ReLU function).

W is the weight matrix applied to the current input x_t.

x_t is the weight matrix applied to the current input W

U is the weight matrix applied to the previous hidden state h_{t-1} .

h_{t-1} is the hidden state at the previous time step.

b is the bias term.

4.2.4 | Long short-term memory networks for depression monitoring

Purpose: LSTM networks, a variation of RNNs, aim to capture long-term dependencies more effectively and prevent the vanishing gradient problem experienced by conventional RNNs.

I. Algorithm: data preprocessing and cleaning: similar to RNNs, LSTM networks require adequate data cleaning, preprocessing, and tokenization. However, they encompass more complex memory units that explicitly manage the flow of information in each cell.

II. Feature extraction: LSTMs inherently manage and extract sequential dependencies. They consist of memory cells, input gates, forget gates and output gates to regulate information flow, making them effective for understanding sequential text data.

III. Training: LSTMs are trained via backpropagation, similarly to RNNs, but with the advantage of the cell state, which regulates the information flow, avoiding the vanishing gradient problem in RNNs.

IV. Classification: similar to RNNs, the trained LSTM model predicts sentiment by considering the long-term dependencies in the input text data.

V. Advantages: LSTMs are highly effective at capturing long-range dependencies in sequential data. They overcome the vanishing gradient problem by their inherent architecture, leading to better performance in learning sequential patterns.

VI. Disadvantages: despite overcoming some RNN limitations, LSTMs may struggle to retain extremely long-term dependencies and can be computationally expensive.

VII. Mathematical equation: the equations in LSTM are more intricate due to their different gates. The mathematical equations for LSTM include the following:

VIII. Forget gate: it determines how much of the previous cell state to retain. It's calculated by combining the previous hidden state and the current input.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f).$$

Input gate: decides how much new information should be added to the cell state. It's computed by combining the previous hidden state and the current input.

$$i_t = \sigma(W_i \ [h_{t-1}, x_t] + b_i).$$

Cell state: computes a new candidate cell state using the hyperbolic tangent function and a weighted sum of the previous hidden state and the current input.

$$g_t = \tanh\left(W_g \ [h_{t-1}, x_t] + \ b_g\right).$$

Cell state update: updates the cell state by integrating the calculations from the forget and input gates.

$$C_t = \ f_t \ C_{t-1} + \ i_t \ g_t.$$

Output gate: decides what the new hidden state should be. It's determined by combining the previous hidden state and the current input.

$$o_t = \ \sigma(W_o \ [h_{t-1}, x_t] + \ b_o).$$

Hidden state output: it computes the new hidden state using the hyperbolic tangent of the updated cell state modified by the output gate.

$$h_t = \ o_t \ \tanh(C_t).$$

Here, σ denotes the sigmoid function, $'.'$ represents element-wise multiplication, and W and b are weights and biases, respectively.

In the quest for accurate depression detection through SA, hybrid models emerge as a strategic solution to harness the combined advantages of lexicon-based methodologies and machine learning algorithms. By integrating these methods, hybrid models aim to mitigate the individual limitations of each approach and create more robust and precise depression detection tools.

4.3 | Hybrid Approach for Depression Monitoring

The hybrid approach combines the strengths of lexicon-based and machine learning techniques to create a synergistic framework for SA within the context of depression monitoring.

Mechanism

I. Data preprocessing: initial data cleaning and tokenization are used to prepare the textual data for analysis.

II. Lexicon-based analysis: the textual data undergoes SA using a lexicon, such as VADER, to determine the emotional tone of the content. This provides a foundational sentiment assessment.

III. Machine learning integration: sentiment scores obtained from lexicon-based analysis, in addition to other textual features, are integrated into a machine learning model.

IV. Feature extraction: the text is transformed into numerical vectors using methods like BoW, TF-IDF, or word2vec, enhancing compatibility with the machine learning model.

V. Training and classification: the machine learning model learns to classify sentiment using labelled data. Integrating lexicon-based sentiment scores enriches the model's understanding of emotional context.

Formulas

The integration of lexicon-based sentiment scores (LB_score) into the machine learning model can be achieved through a weighted combination:

$$\text{Hybrid}_{\text{Score}} = \ \alpha \times \text{LB}_{\text{score}} + \ (1 \ - \ \alpha) \times \ \text{ML}_{\text{score}},$$

where

LB_score: sentiment score obtained from lexicon-based analysis.

ML_score: sentiment score predicted by the machine learning model.

α: weightage parameter, determining the importance of lexicon-based score in the hybrid approach.

Advantages

The hybrid approach leverages the interpretability of lexicon-based SA while harnessing the predictive power of Machine Learning algorithms. Synergizing these two methods aims to capture a more nuanced understanding of sentiment expressions within the context of depression monitoring.

Challenges

Integration challenges arise in combining lexicon-based scores and numerical features for Machine Learning. Additionally, as the lexicon's sentiment assessment might not perfectly align with the true sentiment, discrepancies that impact model performance could emerge. The Hybrid Approach emerges as a promising avenue for depression monitoring. It strives to enrich SA with a holistic view of emotional expressions by merging lexicon-based assessments with data-driven insights. This convergence of methodologies holds the potential to enhance the accuracy and depth of identifying individuals at risk of depression, thereby contributing to timely interventions and support.

Flowchart

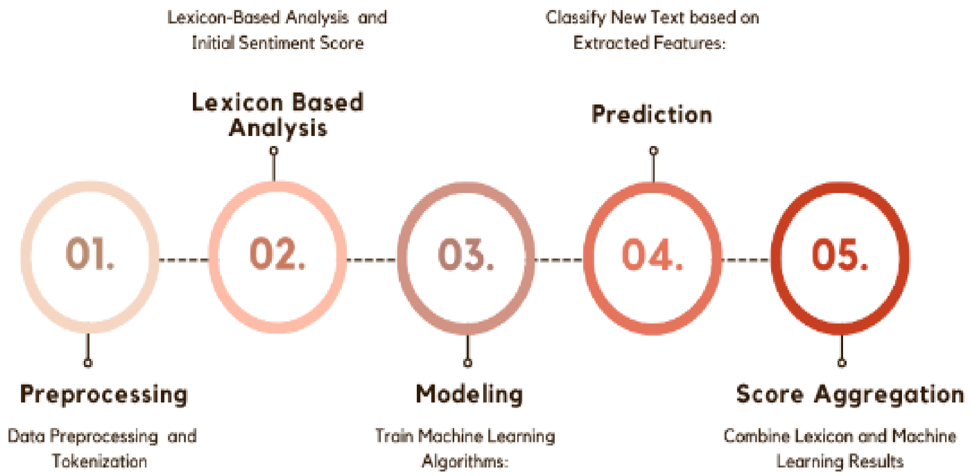

Fig. 4. Flow chart of the hybrid-based model.

5 | Challenges in Sentiment Analysis for Depression Monitoring

Subjectivity in labeling

The inherent subjectivity in labelling depression from text poses a significant challenge. It's challenging to precisely define and label text as depressive or non-depressive due to the broad spectrum of emotional expressions and context sensitivity.

Data quality and diversity

Obtaining diverse and high-quality labelled datasets for training models remains challenging. The scarcity of standardized, labelled datasets encompassing a wide range of demographics, cultures, and linguistic variations hampers the efficacy of SA.

Interpreting context and sarcasm

Context and sarcasm in text present hurdles in SA. Understanding nuanced language use and sarcasm, which might be prevalent in social media or forums, requires sophisticated models to avoid misinterpretations.

Algorithmic bias and generalization

Machine learning models trained on biased data sources might reproduce and perpetuate societal biases in SA, affecting the accuracy and generalizability of these models.

Privacy concerns

Analyzing personal textual data for mental health monitoring raises privacy concerns. Ethical considerations are crucial in handling sensitive data, necessitating stringent anonymization and privacy protection measures.

Domain adaptation

Implementing SA models trained on general text data into the mental health domain requires domain adaptation. Developing models specifically for mental health contexts might enhance accuracy.

6 | Limitations in Existing Research

Imbalance in datasets

Imbalanced datasets, with a higher proportion of non-depressive data compared to depressive data, can lead to biased models with reduced sensitivity in identifying depression.

Model generalization

Generalizing SA models across different platforms or demographics remains a concern. Models trained on specific datasets might lack adaptability to diverse data sources.

Model complexity vs. interpretability

Complex machine learning and deep learning models might offer high accuracy but lack interpretability. The balance between model complexity and interpretability is crucial for practical application.

Lack of real-time analysis

Real-time analysis for immediate support is limited in existing methodologies. Developing models that can provide timely interventions in real-time scenarios remains a challenge.

Limited multimodal analysis

Focusing primarily on textual data limits the comprehensive understanding of user emotions. Integrating other data forms, such as images, videos, or audio, could significantly enhance SA accuracy.

7 | Future Scope in Sentiment Analysis for Depression Monitoring

Multi-lingual and cross-cultural analysis

Developing SA models that can comprehend various languages and cultural nuances is imperative. This enables the creation of more inclusive models suitable for diverse user groups, facilitating a broader application in global contexts.

Mixed-methods approaches

Combining SA with qualitative analysis methods could offer a deeper understanding of depression's manifestation on social media. Qualitative insights complement quantitative findings, providing a holistic perspective on mental health expression.

Integration of clinical data

Merging SA with clinical data, such as electronic health records or self-reported mental health assessments, can augment the accuracy of depression detection. This hybrid approach can validate and enhance the predictive power of SA.

Ethical considerations and transparency

As SA becomes more integral to mental health monitoring, establishing ethical guidelines for data privacy, informed consent, and algorithmic bias is crucial. Ensuring transparent practices helps build trust and reliability.

Longitudinal analysis

Investigating the changes in sentiment over time can assist in identifying early signs of depression or shifts in emotional states. Longitudinal studies provide a deeper understanding of the temporal dynamics of depression expression.

User-centric design

Designing SA tools that cater to users' needs and preferences can improve engagement and accuracy. Allowing users to provide context for their emotions could significantly enhance the outcome of SA.

Interdisciplinary collaboration

Collaborating across disciplines such as computer science, psychology, mental health, and ethics—can lead to more comprehensive SA methodologies. These inclusive efforts can consider various facets of depression detection.

Advancements in algorithm design

Continual advancements in algorithms, specifically tailored for SA in mental health, are essential. Innovations focusing on accuracy, interpretability, and ethical compliance will pave the way for improved tools.

Live monitoring via search history

Integrating search history analysis as part of live monitoring can offer real-time insights into an individual's emotional state. Understanding the sentiment behind their search queries aids in immediate assessment and support, enabling timely interventions.

Chatbots and AI for patient support

Employing a diverse range of chatbots and AI-driven solutions tailored for mental health support is crucial. These platforms can provide personalized assistance, therapeutic guidance, and immediate responses based on SA of user input, thereby enhancing mental health care accessibility and effectiveness.

Integrated wearable technology for emotional analysis

Introducing wearable devices that monitor physiological indicators alongside SA. This integrated approach can provide a more holistic understanding of an individual's emotional state by combining text-generated sentiment with real-time physiological data such as heart rate variability or skin conductance. It enables a more comprehensive assessment of mental health conditions, potentially offering a deeper understanding of emotional fluctuations and stress levels.

8 | Conclusion

This study extensively delved into SA as a potent tool for detecting and monitoring depression through social media and text data. The exploration outlined the varied methods, starting from lexicon-based approaches like VADER and moving on to machine learning models, including Naive Bayes, SVM, and deep learning methods like RNN and Long Short-Term Memory models. Each method offered distinctive advantages and challenges. Moreover, the hybrid approach, synergizing lexicon-based and machine-learning methods, revealed promising prospects.

Throughout this exploration, the study underscored crucial challenges in the landscape, encompassing subjectivity in labelling, data quality, interpretational complexities, privacy concerns, algorithmic bias, and data imbalance. Furthermore, it addressed the limitations concerning model generalization, complexity vs. interpretability trade-offs, and the lack of real-time analysis for immediate interventions.

Looking forward, the paper advocated for multi-lingual and cross-cultural analysis, integration of clinical data, ethical transparency, longitudinal studies, user-centric design, and interdisciplinary collaboration to advance SA methodologies. Additionally, the study envisioned innovative approaches like tracking search history for

live monitoring and employing a wide array of chatbots and AI to enhance patient support systems. In conclusion, while SA stands as a promising avenue for depression detection, it necessitates continual refinement and ethical handling. By embracing technological advancements and interdisciplinary cooperation, the future holds promising opportunities for leveraging SA to improve mental health care, address societal challenges, and offer more effective and accessible support to those affected by depression.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

All data are included in the text.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Reference

[2] Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J., & Qiao, S. (2021). Attention-emotion-enhanced convolutional LSTM for sentiment analysis. *IEEE transactions on neural networks and learning systems*, *33*(9), 4332–4345. https://doi.org/10.1109/TNNLS.2021.3056664

[3] Basiri, M. E., Nemati, S., Abdar, M., Asadi, S., & Acharrya, U. R. (2021). A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. *Knowledge-based systems*, *228*, 107242. https://doi.org/10.1016/j.knosys.2021.107242

[4] Gandhi, U. D., Malarvizhi Kumar, P., Chandra Babu, G., & Karthick, G. (2021). Sentiment analysis on twitter data by using convolutional neural network (CNN) and long short term memory (LSTM). *Wireless personal communications*, 1–10. https://doi.org/10.1007/s11277-021-08580-3

[5] Jain, P. K., Saravanan, V., & Pamula, R. (2021). A hybrid CNN-LSTM: A deep learning approach for consumer sentiment analysis using qualitative user-generated contents. *Transactions on Asian and low-resource language information processing*, *20*(5), 1–15. https://doi.org/10.1145/3457206

[6] Kaur, H., Ahsaan, S. U., Alankar, B., & Chang, V. (2021). A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. *Information Systems Frontiers*, *23*(6), 1417–1429. https://doi.org/10.1007/s10796-021-10135-7

[7] Pathak, A. R., Pandey, M., & Rautaray, S. (2021). Topic-level sentiment analysis of social media data using deep learning. *Applied soft computing*, *108*, 107440. https://doi.org/10.1016/j.asoc.2021.107440

[8] Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future generation computer systems*, *115*, 279–294. https://doi.org/10.1016/j.future.2020.08.005

[9] Wang, K., Shen, W., Yang, Y., Quan, X., & Wang, R. (2020). *Relational graph attention network for aspect-based sentiment analysis*. https://doi.org/10.48550/arXiv.2004.12362

[10] Zhang, T., Gong, X., & Chen, C. L. P. (2021). BMT-Net: broad multitask transformer network for sentiment analysis. *IEEE transactions on cybernetics*, *52*(7), 6232–6243. https://doi.org/10.1109/TCYB.2021.3050508

[11] Dashtipour, K., Ieracitano, C., Carlo Morabito, F., Raza, A., & Hussain, A. (2021). An ensemble based classification approach for persian sentiment analysis. In *Progresses in artificial intelligence and neural systems* (pp. 207–215). Springer. https://doi.org/10.1007/978-981-15-5093-5_20

[12] Sabri, N. M., & Mohamad, N. A. (2022). Detection of suicidal tweets based on naive bayes algorithm. *International journal of advanced research in technology and innovation*, *4*(3), 47–59. https://doi.org/10.55057/ijarti.2022.4.3.6

[13] Samareh, A., Jin, Y., Wang, Z., Chang, X., & Huang, S. (2018). Detect depression from communication: How computer vision, signal processing, and sentiment analysis join forces. *IISE transactions on healthcare systems engineering*, *8*(3), 196–208. https://doi.org/10.1080/24725579.2018.1496494

[14] Sheikh, S., Sharma, M., & Singh, A. (2023). *Recent advances in computing sciences: proceedings of RACS 2022*. CRC Press.

[15] Kim, J. Y., Liu, N., Tan, H.-X., & Chu, C.-H. (2017). Unobtrusive monitoring to detect depression for elderly with chronic illnesses. *IEEE Sensors Journal*, *17*(17), 5694–5704. https://doi.org/10.1109/JSEN.2017.2729594

[16] Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature. https://books.google.com/books?hl=en&lr=lang_en&id=xYhyEAAAQBAJ&oi=fnd&pg=PP1&dq=Liu,+B.+(2022).+Sentiment+analysis+and+opinion+mining.+Springer+Nature.&ots=rlUDKFS2xA&sig=92zi4n_nROWO-SKUVHMUk6tcQhE

[17] Mustafa, R. U., Ashraf, N., Ahmed, F. S., Ferzund, J., Shahzad, B., & Gelbukh, A. (2020). A multiclass depression detection in social media based on sentiment analysis. *17th International Conference on Information Technology–New Generations (ITNG 2020)* (pp. 659-662). Springer International Publishing. https://doi.org/10.1007/978-3-030-43020-7_89

[18] Ushashree, P., Harshika, G., Haani, U., & Kalai, R. (2021). Sentiment analysis to detect depression in social media users: overview and proposed methodology. *Emerging research in computing, information, communication and applications* (pp. 469–480). Springer. https://doi.org/10.1007/978-981-16-1338-8_39

[19] Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia computer science*, *161*, 707–714. https://doi.org/10.1016/j.procs.2019.11.174

[20] Suman, S. K., Shalu, H., Agrawal, L. A., Agrawal, A., & Kadiwala, J. (2020). *A novel sentiment analysis engine for preliminary depression status estimation on social media*. https://doi.org/10.48550/arXiv.2011.14280

[21] Jo, H., Kim, S.-M., & Ryu, J. (2017). *What we really want to find by sentiment analysis: the relationship between computational models and psychological state*. https://doi.org/10.48550/arXiv.1704.03407

[22] IIslam, M. R., Kamal, A. R. M., Sultana, N., Islam, R., & Moni, M. A. (2018). Detecting depression using k-nearest neighbors (KNN) classification technique. *2018 international conference on computer, communication, chemical, material and electronic engineering (IC4ME2)* (pp. 1-4). IEEE. https://doi.org/10.1109/IC4ME2.2018.8465641

[23] Li, A., Jiao, D., & Zhu, T. (2018). Detecting depression stigma on social media: A linguistic analysis. *Journal of Affective Disorders*, *232*, 358–362. https://doi.org/10.1016/j.jad.2018.02.087